**EVALUACIÓN DE LOS RESULTADOS**

16

**Recuperación de Información**

**Evaluación**

**Jorge Andrés Galindo 679155 Javier Aranda García 679184**

Tras realizar el proceso de evaluación de los resultados con el programa desarrollado, se ha procedido a mirar la tabla para saber a qué letra correspondía nuestro grupo. Observando lo resultados, todas las medidas coincidían con las expuestas para el grupo “i”, tanto las medidas generales, como los obtenidos de los valores precisión-recall interpolados.

Para los valores que hemos obtenido, que comparándolos con el resto de los grupos son de los mejores, creemos que han influido mucho distintos aspectos que tuvimos en cuenta en el proceso de diseño. En primer lugar, la separación de campos que realizamos era la obvia, pero decidimos juntar en un mismo campo la descripción del trabajo con el creador del mismo. Esto fue debido a que encontrar patrones para detectar nombres propios o apellidos no era obvio, y podían perjudicar más al sistema que ayudarlo (confundir un país como apellido, el nombre de un río…).

Otro aspecto que tuvimos en cuenta para el diseño, fue la utilización de un analizador español eliminando la indexación de descripciones en inglés de trabajos. Esta decisión se tomó porque las consultas iban a realizarse en español, y tener una descripción en inglés, unida a la que había en español, podría bajar considerablemente el “score” de una consulta. Además, se creó un campo extra para fechas contenidas en los campos descripción y título de los documentos, lo que unido al ya fijado campo fecha, podía presentar una gran mejora en resultados relacionados con años.

Además, se utilizó un modelo de tratamiento probabilístico; es decir, se eliminó el utilizado en prácticas basado en vectores, ya que se estuvieron comparando los resultados de ambos, y daba mejores resultados el modelo basado en probabilidad. Finalmente, un factor que también pudo tener que ver con los resultados, fue el amplio tratamiento que se dio a la búsqueda de patrones basados en fechas en las necesidades de información. Se introdujeron rangos para fechas concretas, años anteriores, posteriores…

En nuestra opinión, si tuviéramos que basarnos en una medida para reflejar la calidad de un sistema de recuperación, en primer lugar, miraríamos la precisión a un nivel concreto de recuperación. Esta medida es muy útil ya que evalúa la precisión de los “k” primeros documentos devueltos (medida válida para búsquedas web que solo se miran las dos primeras páginas de resultados). La otra medida útil sería la medida de precisión-recall interpolada, ya que un usuario que está realizando una búsqueda está dispuesto a ver más resultados si estos son relevantes para su necesidad.

Viendo los resultados de todos los grupos, y dejando al margen nuestro sistema de recuperación, seleccionaríamos sin duda el grupo “b”. Esto es debido a que los valores de precisión y recall son bastante elevados (rondando el 50% ambos), y la precisión de los 10 primeros resultados es altísima (60%), lo que da a entender que los documentos devueltos en las primeras posiciones son de un alto interés al sujeto que formula la búsqueda. Además, la curva precisión-recall interpolada muestra que, para un recall bajo tiene una alta precisión, pero a medida que aumenta el recall es el sistema que mejor mantiene una cierta precisión.